

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



①

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> <b>C12N 15/12, C07K 15/06, 15/14</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 91/15583</b> <b>(43) International Publication Date:</b> 17 October 1991 (17.10.91)
<b>(21) International Application Number:</b> PCT/US91/02344 <b>(22) International Filing Date:</b> 4 April 1991 (04.04.91) <b>(30) Priority data:</b> 504,800 5 April 1990 (05.04.90) US <b>(71) Applicant:</b> THE AMERICAN NATIONAL RED CROSS [US/US]; 430 17th Street, N.W., Washington, DC 20006 (US). <b>(72) Inventors:</b> HLA, Timothy, Tun ; 20424 Cabana Drive, Germantown, MD 20876 (US). MACIAG, Thomas ; 6050 Valerian Lane, Rockville, MD 20852 (US). <b>(74) Agents:</b> GOLDSTEIN, Jorge, A. et al.; Sterne, Kessler, Goldstein & Fox, 1225 Connecticut Avenue, N.W., Suite 300, Washington, DC 20036 (US).		<b>(81) Designated States:</b> AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL (European patent), NO, PL, RO, SD, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> A PROTEIN FAMILY RELATED TO IMMEDIATE-EARLY PROTEIN EXPRESSED BY HUMAN ENDOTHELIAL CELLS DURING DIFFERENTIATION  <b>(57) Abstract</b> <p>This invention provides a novel family of tissue specific genes and proteins that are related to a G-protein-coupled receptor gene and the receptor protein. The gene is an intermediate early gene that is expressed in differentiating endothelial cells. In particular, this invention provides a gene, <i>edg-1</i>, that is an immediate-early gene that encodes a G-protein-coupled receptor in endothelial cells. This invention also provides the G-protein-coupled receptor protein that is encoded by <i>edg-1</i>.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TC	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark				

A PROTEIN FAMILY RELATED TO IMMEDIATE-EARLY PROTEIN EXPRESSED  
BY HUMAN ENDOTHELIAL CELLS DURING DIFFERENTIATION

1  
2  
3

4 BACKGROUND OF INVENTION

5 The endothelium is composed of a monolayer of quiescent  
6 cells, endothelial cells. Endothelial cells, which form the  
7 inner lining of blood vessels participate in a multiplicity  
8 of physiological functions, including the formation of a  
9 selective barrier for the translocation of blood constituents  
10 and macromolecules to underlying tissues and the maintenance  
11 of a non-thrombogenic interface between blood and tissue.  
12 Endothelial cells are also an important component in the  
13 development of new capillaries and blood vessels. Blood  
14 vessel development, which is called angiogenesis, occurs  
15 during developmental periods, such as during development of  
16 the vascular system, and as part of the pathophysiology of a  
17 variety of disease states, such as psoriasis, arthritis,  
18 chronic inflammatory conditions, diabetic retinopathy, and  
19 tumor development.

20 Angiogenesis, which involves the organized migration,  
21 proliferation, and differentiation of the endothelial cells,  
22 is initiated by the endothelial cell in response to angiogenic  
23 stimuli and can be separated into three distinct events: cell  
24 migration, cell proliferation and cell differentiation,  
25 whereby the cells organize into a tubular structure.

26 These events are mediated in vitro, and most likely in  
27 vivo, by mitogenic polypeptides. The migration of endothelial  
28 cells is induced by factors, including the heparin binding

-2-

1 growth factors and angiotropin. Proliferation is induced by  
2 the heparin binding growth factors (hereinafter HBGFs) and  
3 differentiation and cellular organization is induced by  
4 polypeptides, including interleukin-1 (hereinafter IL-1),  
5 tumor necrosis factor (hereinafter TNF), gamma-interferon,  
6 transforming growth factor alpha and beta (hereinafter TGF- $\alpha$   
7 and TGF- $\beta$ , respectively) and phorbol myristic acetate  
8 (hereinafter PMA).

9 The extracellular matrix (hereinafter ECM), which  
10 contains numerous components, also modulates endothelial cell  
11 differentiation. If endothelial cells are cultured in vitro  
12 on collagen gels in the presence of PMA organized networks of  
13 tubular structures form, and, if the cells are cultured in ECM  
14 conditioned medium the formation of tubular structures is  
15 accelerated.

16 The importance of the ECM components for mediation of  
17 endothelial cell differentiation is evidenced by the  
18 observations that antibodies that have been prepared against  
19 fibronectin and laminin inhibit formation of the  
20 differentiated phenotype, while proteolytic modification of  
21 fibronectin by plasmin leads to rapid modification of the  
22 endothelial cell phenotypic changes that are observed in  
23 vitro. In addition, competitive inhibitors of the laminin  
24 and fibronectin receptor binding domains also inhibit the  
25 ability of endothelial cells to complete the non-terminal  
26 differentiation program.

27 As discussed above, the polypeptide cytokines and PMA  
28 inhibit the HBGF-1-induced proliferation of endothelial cells  
29 and induce differentiation thereof. These factors induce a  
30 reversible phenotypic transition from a non-polar cobblestone  
31 monolayer into a polar elongated, fibroblast-like phenotype.  
32 The inhibition of HBGF-1-induced proliferation is mediated,  
33 at least in part, via down regulation of the HBGF-1 receptor.

-3-

1           It is also known that PMA activates protein kinase C,  
2           which a family of phospholipid- and calcium-activated protein  
3           kinases. This activation results in the transcription of an  
4           array of proto-oncogene transcription factors, including c-  
5           fos, c-myc and c-jun, proteases, protease inhibitors,  
6           including collagenase type I and plasminogen activator  
7           inhibitor, and adhesion molecules, including intercellular  
8           adhesion molecule I. Protein kinase C activation antagonizes  
9           growth factor activity by the rapid phosphorylation of the  
10          epidermal growth factor receptor. Phosphorylation decreases  
11          tyrosine kinase activity.

12          Upon induction of differentiation of endothelial cells  
13          in vitro by a cytokine or PMA, a set of immediate-early genes  
14          are rapidly induced via a pathway that does not require  
15          protein synthesis. Included among these immediate-early genes  
16          are transcriptional factors, cytokines, cytoskeletal proteins,  
17          nuclear hormone receptors and extracellular matrix receptors.

18          Cell surface receptors bind circulating signal  
19          polypeptides, such as growth factors and hormones, as the  
20          initiating step in the induction of numerous intracellular  
21          effector functions. Receptors are classified on the basis of  
22          the particular type of pathway that is induced. Included  
23          among these classes of receptors are those that bind growth  
24          factors and have intrinsic tyrosine kinase activity, such as  
25          the HBGF receptors and those that couple to effector proteins  
26          through guanine nucleotide binding regulatory proteins,  
27          hereinafter referred to as G-protein coupled receptors and G-  
28          proteins, respectively. The G-protein transmembrane signaling  
29          pathways consist of three proteins: receptors, G proteins and  
30          effectors.

31          G proteins, which are the intermediaries in transmembrane  
32          signaling pathways, are heterodimers and consist of  $\alpha$ ,  $\beta$  and  
33          gamma subunits. Among the members of a family of G proteins

1 the  $\alpha$  subunits differ. Functions of G proteins are regulated  
2 by the cyclic association of GTP with the  $\alpha$  subunit followed  
3 by hydrolysis of GTP to GDP and dissociation of GDP.

4 G-protein coupled receptors are a diverse class of  
5 receptors that mediate signal transduction by binding to G-  
6 proteins. Signal transduction is initiated via ligand binding  
7 to the cell membrane receptor, which stimulates binding of the  
8 receptor to the G-protein. The receptor-G-protein interaction  
9 releases GDP, which is specifically bound to the G-protein,  
10 and permits the binding of GTP, which activates the G-protein.  
11 Activated G-protein dissociates from the receptor and  
12 activates the effector protein, which regulates the  
13 intracellular levels of specific second messengers. Examples  
14 of such effector proteins include adenylyl cyclase, guanylyl  
15 cyclase, phospholipase C, and others.

16 G-protein-coupled receptors, which are glycoproteins, are  
17 known to share certain structural similarities and homologies  
18 (see, e.g., Gilman, A.G., Ann. Rev. Biochem. 56: 615-649  
19 (1987), Strader, C.D. et al. The FASEB Journal 3: 1825-1832  
20 (1989), Kobilka, B.K., et al. Nature 329: 75-79 (1985) and  
21 Young et al. Cell 45: 711-719 (1986)). Among the G-protein-  
22 coupled receptors that have been identified and cloned are the  
23 substance K receptor, the angiotensin receptor, the  $\alpha$ - and  $\beta$ -  
24 adrenergic receptors and the serotonin receptors. G-protein-  
25 coupled receptors share a conserved structural motif. The  
26 general and common structural features of the G-protein-  
27 coupled receptors are the existence of seven hydrophobic  
28 stretches of about 20-25 amino acids each surrounded by eight  
29 hydrophilic regions of variable length. It has been  
30 postulated that each of the seven hydrophobic regions forms  
31 a transmembrane  $\alpha$  helix and the intervening hydrophilic  
32 regions form alternately intracellularly and extracellularly

-5-

1 exposed loops. The third cytosolic loop between transmembrane  
2 domains five and six is the intracellular domain responsible  
3 for the interaction with G-protein.

4 G-protein-coupled receptors are known to be inducible.  
5 This inducibility was originally described in lower  
6 eukaryotes. For example, the cAMP receptor of the cellular  
7 slime mold, Dictyostelium, is induced during differentiation  
8 (Klein et al., Science 241: 1467-1472 (1988)). During the  
9 Dictyostelium discoideum differentiation pathway, cAMP,  
10 induces high level expression of its G-protein-coupled  
11 receptor. This receptor transduces the signal to induce the  
12 expression of the other genes involved in chemotaxis, which  
13 permits multicellular aggregates to align, organize and form  
14 stalks (see, Firtel, R.A., et al. Cell 58: 235-239 (1989) and  
15 Devreotes, P., Science 245: 1054-1058 (1989)). H u m a n  
16 endothelial cells utilize a series of morphological correlates  
17 during its differentiation pathway, discussed supra., in which  
18 individual cells migrate, align and organize to form  
19 multicellular capillary-like structures.

#### 20 SUMMARY OF THE INVENTION

21 It is one object of this invention to provide a novel G-  
22 protein-coupled receptor that is the product of an immediate  
23 early gene that is expressed in endothelial cells during the  
24 early stage of differentiation.

25 It is another object of this invention to provide a  
26 family of proteins that are expressed in a tissue-specific  
27 manner and that are related to the novel G-protein-coupled  
28 receptor that is the product of an immediate early gene that  
29 is expressed in endothelial cells during the early stage of  
30 differentiation.

1           It is another object of this invention to provide DNA  
2 molecules that encode each member of the family of proteins  
3 that are expressed in a tissue-specific manner and that are  
4 related to the novel G-protein-coupled receptor that is the  
5 product of an immediate early gene that is expressed in  
6 endothelial cells during the early stage of differentiation.

7           It is another object of this invention to provide DNA  
8 molecules that encode the novel G-protein-coupled receptor  
9 that is the product of an immediate early gene that is  
10 expressed in endothelial cells during the early stage of  
11 differentiation.

12           In accordance with this invention there is provided a DNA  
13 molecule that encodes edg-1 gene product, which is the product  
14 of an immediate-early gene that is expressed in the early  
15 stage of differentiation of endothelial cells in response to  
16 PMA or IL-1.

17           This invention provides a gene and protein, which is the  
18 first immediate-early gene that encodes a G-protein-coupled  
19 receptor.

20           Unless defined otherwise, all technical and scientific  
21 terms used herein have the same meaning as is commonly  
22 understood by one of ordinary skill in the art to which this  
23 invention belongs. Although methods and materials similar or  
24 equivalent to those described herein can be used in the  
25 practice of testing of the present invention, the preferred  
26 methods and materials are now described. All publications  
27 mentioned hereunder are incorporated by reference.



-7-

1                    BRIEF DESCRIPTION OF THE FIGURES

2            Figure 1. The identification of edg-1, an Immediate  
3            early gene induced by PMA in HUVEC (human umbilical vein  
4            endothelial cells).

5            Confluent cultures of HUVEC were treated with 20 ng/ml  
6            of PMA for the indicated times. The cells were then lysed,  
7            RNA purified, and total RNA (10 µg) analyzed by Northern blot  
8            analysis. The cDNA probes that were used were edg-1 (A) and  
9            glyceraldehyde-3-phosphate (GAPDH) (B) cDNA.

10           Figure 2. Confluent cultures of HUVEC were treated with  
11           the indicated reagents for 4 hour and the RNA was isolated.  
12           Total RNA (10 µg) was fractionated by 1% agarose-formaldehyde  
13           gel electrophoresis, blotted onto a zeta-probe membrane and  
14           hybridized with [<sup>32</sup>P]-labeled edg-1 (A) or a GAPDH (B) cDNA  
15           probes. The following reagents were used: PMA (20 ng/ml), chx  
16           (5 µg/ml), Actinomycin D (Act D) (2 µg/ml). Each reagent was  
17           used either alone or in combination.

18           Figure 3. Confluent cultures of HUVEC were pre-treated  
19           with 20 ng/ml PMA for 4 hour. Either Act D (2 µg/ml) alone or  
20           with chx (5 µg/ml) was added to the cultures, at a time  
21           designated 0. At the indicated time points, cultures were  
22           harvested and Northern blot analysis was performed on total  
23           RNA as described above using the edg-1 (A) and GAPDH (B) cDNA  
24           probes.

25           Figure 4. HUVEC were either untreated or treated with  
26           20 ng/ml PMA for 2 hour after which nuclei were prepared.  
27           Run-off transcripts were obtained by labelling 10<sup>7</sup> nuclei in  
28           vitro with [<sup>32</sup>P]-UTP. RNA was purified and hybridized to  
29           immobilized plasmid DNA encoding edg-1 (10 µg/slot), human  
30           fibronectin (fn) (2 µg/slot) and pBluescript (pBS) (10  
31           µg/slot).

1        Figure 5. Nucleotide and Deduced Amino Acid Sequence of  
2 Human edg-1.

3        The nucleotide (1-2774) and deduced amino acid sequence  
4 (1-380) is shown for human edg-1 cDNA. The deduced  
5 transmembrane domains are underline and potential N-linked  
6 glycosylation sites are shown with ann asterisk. Possible  
7 serine and threonine phosphorylation sites are shown with  
8 closed circles. The basic amino acid-rich intracellular  
9 domain, which is located between transmembrane domains five  
10 and six is highlighted with open circles. The Kozak consensus  
11 translation initiation sequence (5') and polyadenylation sites  
12 (3') are shown with double lines underneath their respective  
13 sequences. The Genbank accession number for this nucleotide  
14 sequence is M31210.

15        Figure 6. The amino acid sequence of the putative edg-  
16 1 translation product was aligned with Substance K receptor  
17 (SKR), Substance P receptor (SPR),  $\beta_2$ -adrenergic receptor  
18 (B2AR), Serotonin receptor 1c (5HTC),  $\alpha_2$ -adrenergic receptor  
19 (A2A), Serotonin receptor 1a (5HT1a), Rhodopsin (OSPD) and  
20 angiotensin receptor (MAS). Highly homologous regions are  
21 boxed and indicated on the linear schematic.

22        Figure 7. A structural model for the putative edg-1  
23 translation product is shown. This model is analogous to other  
24 G-protein-coupled receptors. The potential N-linked  
25 glycosylation sites are indicated with an inverted "Y".  
26 Potential phosphorylation sites at serine and threonine  
27 residues are shown with dark circles. The third cytosolic  
28 intracellular domain, which is between transmembrane domains  
29 5 and 6 contains a highly basic region (11/35 residues) is  
30 also indicated.

1           Figure 8. Hydrophobicity Profile of edg-1 Translation  
2 Product. The deduced amino acid sequence of edg-1 was  
3 analyzed for hydrophobic regions and the amino acid sequence  
4 (residues) plotted against the hydrophobicity index. The  
5 putative transmembrane (TM) domains are indicated.

6           Figure 9. Expression of edg-1 transcript in human cells.  
7 Total RNA (5  $\mu$ g) from human saphenous vein smooth muscle  
8 cells (S), foreskin fibroblasts (F), HeLa cells (H),  
9 epidermoid carcinoma (A431) cells (A), melanocytes (M), brain  
10 tissue (B) and endothelial cells (E) were reverse transcribed  
11 into cDNA and amplified with edg-1 specific oligonucleotide  
12 primers that span the carboxy-terminal tail domain (A) and  
13 the third cytosolic loop (B). Amplified DNA was separated  
14 by agarose gel electrophoresis and visualized by ethidium  
15 bromide staining. Molecular weight markers (indicated by  
16 arrows) are from top to bottom: 1.6 Kb, 1.0 Kb, 0.5 Kb, 0.4  
17 Kb, 0.3 Kb, 0.2 Kb and 0.15 Kb.

18           It can be seen in (A) that transcript of the expected  
19 size, about 600 bp,, which was amplified using oligonucleotide  
20 primers specific for the C-terminal domain, was present in  
21 RNA from all the cultured cell lines and human brain. In  
22 contrast, when the transcript was amplified using an a pair  
23 oligonucleotides that span the third intracellular loop, cell  
24 or tissue specific bands were observed.

#### 25           DESCRIPTION OF THE PREFERRED EMBODIMENTS

26           In the invention described herein a novel gene, edg, and  
27 the protein encoded thereby has been identified. In addition,  
28 this invention provides a family of proteins that are  
29 structurally and functionally related to this protein as well  
30 as DNA molecules, but that are tissue or cell type specific  
31 are provided.

-10-

1       As used her in, the edg-G-protein-c upled receptor family  
2       is a family of related proteins that share substantial  
3       homology and structure and that contain common constant  
4       regions or domains but differ in at least one variable region  
5       or domain that includes the third cytosolic loop. See, e.g.,  
6       Figures 6, 7, and 9. The particular variable region and,  
7       thus, each family member, is expressed in a tissue-specific  
8       manner.

9       As used herein, expression of a transcript in a tissue-  
10      specific manner includes expression of transcripts that are  
11      expressed in only certain tissues or cell types. Such tissue-  
12      specific expression can be effected through a variety of  
13      mechanisms, including the expression of different genes in  
14      each tissue or cell type, through alternative splicing of the  
15      same gene in each tissue or cell type, or through  
16      recombination of germ line DNA in during development or  
17      differentiation of each cell type.

18      As used herein, the edg-1-G-coupled protein receptor  
19      transcript is the intermediate early transcript that is  
20      expressed in the early stage of differentiation in endothelial  
21      cells that can be induced or stimulated with PMA and  
22      interleukin-1 (IL-1) but not with TGF- $\beta$ , HBGF-1, or  $\alpha$ -  
23      thrombin. The edg-1 G-coupled protein receptor transcript  
24      encodes the edg-1 G-coupled protein receptor.

25      As used herein, the edg-1-G-coupled protein receptor  
26      transcript family is a family of transcripts that are  
27      expressed in a tissue-specific manner and encode members of  
28      the family of related proteins that share substantial homology  
29      and structure and that contain common constant regions or  
30      domains but differ in at least one variable region that  
31      includes the third cytosolic loop.

-11-

1           As used herein, DNA encoding a protein includes any DNA  
2 molecule that encodes a protein that has substantially the  
3 same amino acid sequence. Each of such proteins may, however,  
4 differ at sites that are not essential to protein function and  
5 includes proteins isolated from different individuals in the  
6 same species, proteins isolated from different species that  
7 share substantially the same biological activities, and  
8 proteins isolated from different cultured cell lines.

9           As used herein, the *edg-1* transcript refers to the 2.8  
10 Kb (about 3 Kb) transcript that encodes the receptor protein.  
11 This term is herein used interchangeably with the *edg*  
12 transcript, *edg* mRNA. The *edg-1* transcript also refers to this  
13 transcript, but also refers to the 1-Kb clone that was  
14 isolated from the differential screen, which contained a poly  
15 A tract at 3' end, a unique nucleotide sequence and hybridized  
16 to the about 3.0 Kb PMA inducible mRNA species, the *edg-1*  
17 transcript.

18           Because PMA inhibits endothelial cell proliferation and  
19 induces differentiation, the identification and isolation of  
20 immediate-early genes yields insight into the molecular  
21 mechanisms involved in the regulation of endothelial cell  
22 differentiation.

23           Immediate-early genes that are expressed in endothelial  
24 cells may be isolated from any source of endothelial RNA. In  
25 one embodiment of this invention, human umbilical vein  
26 endothelial cells (hereinafter HUVEC) are used. The HUVEC are  
27 either untreated and treated with PMA, IL-2 or any other  
28 signal that induces these genes.

-12-

1       The desired immediate-early genes can be identified by  
2       any means in which the transcripts comparing the transcripts  
3       in cells that are stimulated with PMA, IL-2 or other inducer  
4       with the transcripts that are present in untreated cells.  
5       Those that are present only in the treated cells are, thus,  
6       immediate-early genes. In addition, any member of the G-  
7       protein-coupled receptor family of this invention can be  
8       identified by screening an appropriate library with an  
9       appropriate probe derived from the *edg-1* clone. For example,  
10      an appropriate probe would be one derived from the 3' end of  
11      the clone. Any methods known to those of skill in the art to  
12      accomplish this may be used.

13      In endothelial cells the immediate-early gene of this  
14      invention is the *edg-1* encoding gene. It is induced by IL-1,  
15      LPS or PMA, but not by HBGF-1, TGF- $\beta$ , or  $\alpha$ -thrombin. The *edg-1*  
16      clone provided herein encodes a protein that shares many  
17      structural and sequence similarities with known G-protein-  
18      coupled receptors, including the  $\beta$ -adrenergic, substance K,  
19      substance P, rhodopsin, serotonin (5-HT), tachykinin receptors  
20      and the cAMP receptor of Dictyostelium.

21      The N-linked glycosylation site at Asn<sub>30</sub> is also found in  
22      the Substance K and angiotensin receptors. The two N-linked  
23      glycosylation sites are found within the amino-terminal domain  
24      of all G-protein-coupled receptors. The region in proximity  
25      to the second and third hydrophobic domains is highly  
26      conserved among all such receptors, including that encoded by  
27      *edg-1*. In the  $\beta_2$ -adrenergic receptor Asp<sub>130</sub> is known to be  
28      absolutely necessary for G-protein; in the *edg-1*-encoded  
29      protein the Asp/Glu-Arg is conserved.

30      Although the overall sequence similarity between the  
31      *edg-1* G-protein-coupled receptor of this invention and other  
32      such receptor is quite divergent, there is a significant  
33      degree of sequence similarity within the carboxy-terminal

-13-

1 half, particularly within transmembrane domain seven. It is  
2 most similar to those receptors that recognize peptides as  
3 receptor ligands.

4 The intracellular hydrophilic loop regions contain four  
5 potential phosphorylation sites at residues Thr<sub>72</sub>, Ser<sub>231</sub>, Thr<sub>235</sub>  
6 and at Ser<sub>351</sub>. This feature is common to many G-protein-  
7 coupled receptors. Phosphorylation at the Ser and Thr  
8 residues within the intracellular domains has been implicated  
9 in the phenomenon of receptor desensitization.

10 The hydrophilic region between transmembrane domains five  
11 and six is the region that is absolutely necessary for G-  
12 protein coupling and it is highly divergent among members of  
13 the G-protein-coupled receptor proteins. In the G-protein-  
14 coupled receptor that is encoded by edg-1, this region is  
15 highly basic. The family of edg-1 related tissue-specific  
16 proteins provided in this invention differ in this region and,  
17 thus, most likely differ in their respective binding or  
18 coupling interactions with the G-protein or protein ligands.

19 The ligand that binds to each of the members of the  
20 family of G-protein-coupled receptor proteins of this  
21 invention can be identified by methods that are known to those  
22 of skill in the art. For example, xenopus oocytes can be  
23 transfected with DNA that encodes the particular protein. The  
24 protein will be expressed on the cell surface of the oocytes.  
25 Since these oocytes are sensitive to calcium exchange across  
26 the cell membrane, binding of the appropriate ligand causes  
27 calcium exchange across membrane. Labeled calcium can be used  
28 and the ligand that causes labeled calcium exchange can be  
29 identified. Among the candidates for the ligand that binds  
30 to the edg-1-G-protein coupled receptor are ATP, AMP,  
31 adenosine, leukotrienes, prostenoids, histamine, bombasin,  
32 thrombin, azopressin, bradykinin, endothelin, serotensin,  
33 substance P and neuropeptide.

-14-

1           The following examples are included for illustrative  
2 purposes only and are not intended to limit the scope of the  
3 invention.

#### 4                           EXAMPLE 1

##### 5                           Materials and Cell Culture

6           Recombinant human interleukin  $\alpha$  (IL-1 $\alpha$ ), which was the  
7 gift of Dr. Peter Lomedico, Hoffman La Roche, Nutley, NJ.  
8 Recombinant human HBGF-1 $\alpha$  was obtained from Anthony Jackson,  
9 American Red Cross, Rockville, MD. Porcine TGF- $\beta$  was purchased  
10 from R & D Systems.

11           Primary cultures of human umbilical vein endothelial  
12 cells (HUVEC) were obtained from Dr. Michael Gimbrone, Harvard  
13 Medical School, Boston, MA, and were grown on fibronectin-  
14 coated plates in Medium 199 supplemented with 10% (v/v) fetal  
15 bovine serum, 1x antibiotic and antimycotic mixture (GIBCO,  
16 Grand Island, NY), 150  $\mu$ g/ml crude endothelial cell growth  
17 factor (Maciag et al., 1981) and 5 U/ml heparin (Sigma) as  
18 described in Maciag et al. ((1981) J. Biol. Chem. 91, 420-  
19 426). Cells were subcultured at a 1:5 split ratio and  
20 cultures between passages of 4 and 12 were used. At  
21 confluence, cells were maintained in medium without the growth  
22 factor and heparin for two days to achieve quiescence.

##### 23                           RNA Preparation and cDNA Library Construction

24           Total RNA was obtained from cells that either untreated  
25 or treated with 20 ng/ml PMA (Sigma) and 5  $\mu$ g/ml of  
26 cycloheximide (hereinafter chx) (Sigma) for 4 hours. The  
27 cells were rinsed with phosphate-buffered saline, lysed in 4M  
28 guanidinium isothiocyanate and total RNA purified as described  
29 in Winkles, J., et al. ( (1987) Proc. Natl. Acad. Sci. USA 84,



-15-

1 7124-7128). Poly A<sup>+</sup> RNA (10 µg) from HUVEC exposed to PMA  
2 and chx was converted to double-stranded cDNA and cloned into  
3 the Eco RI site of lambda gt10, using the cDNA synthesis kit  
4 from Bethesda Research Labs (Gaithersburg, MD) and the cDNA  
5 cloning kit from Amersham (Chicago, IL). The library contained  
6 > 10<sup>6</sup> independent clones, with an average insert size of  
7 approximately 1 Kb.

8 Northern Blot Analysis.

9 Total RNA (10 µg) was electrophoresed on a 1% agarose  
10 gel containing 2.2 M formaldehyde, capillary-blotted onto  
11 Zeta-probe membrane (Biorad) and UV cross-linked (Maniatis et  
12 al. (1982) In Molecular Cloning: A Laboratory Manual, Cold  
13 Spring Harbor Laboratory, Cold Spring Harbor, NY). The cDNA  
14 insert fragment for edg-1 (2.8 Kb) or human GAPDH (1 Kb) was  
15 labeled to high specific activity (>10<sup>8</sup> cpm/µg) using a random  
16 primer labeling kit (BRL) and was used to hybridize filters  
17 in Church-Gilbert buffer (0.5 M sodium phosphate pH 7.2,  
18 containing 7% SDS and 1% bovine serum albumin, 1mM EDTA and  
19 20% formamide at 65° C for 16-20 hrs. Filters were washed  
20 twice for 15 min at high-stringency (0.1xSSC, 65° C).

21 Differential Screening of cDNA Library

22 The differential screen was performed by plating 2 x 10<sup>6</sup>  
23 pfu of the library onto bacteriological plaques (15 cm  
24 diameter) containing LB agar. The phage were allowed to grow  
25 at 37° C until plaques were approximately 0.5 mm in diameter.  
26 Phage DNA was adsorbed onto Gene-screen plus nylon filters  
27 (Dupont, DE), in duplicate, denatured, neutralized, and UV  
28 cross-linked.

-16-

1       The probe for differential screening was prepared by  
2 reverse transcription of 1  $\mu$ g of poly A<sup>+</sup> RNA from control and  
3 PMA/chx-treated HUVEC. The reaction conditions were as  
4 follows: 50 mM Tris HCl, pH 8.3, 75 mM KCl, 20 mM  
5 dithiothreitol, 3 mM MgCl<sub>2</sub>, 500  $\mu$ Ci [<sup>32</sup>P]- $\alpha$ -dCTP, 20  $\mu$ M dCTP,  
6 200  $\mu$ M each of dATP, dCTP, and dTTP, 0.5  $\mu$ g/ml of oligo dT<sub>12-18</sub>  
7 and 400 units of MMLV-reverse transcriptase (Bethesda  
8 Research Labs, Gaithersburg, MD).

9       After incubation at 37° C for 60 minutes, RNA was  
10 hydrolyzed by treatment with 100  $\mu$ l 0.6M NaOH and 20 mM EDTA  
11 for 30 minutes at 65° C. The cDNA was purified on Sephadex  
12 G-50 columns and ethanol-precipitated. Duplicate filters were  
13 incubated with 10 cpm/ml of cDNA for 48 hours at 65° C in  
14 hybridization buffer containing 2% SDS, 1 M NaCl and 10%  
15 dextran sulfate. The filters were washed twice for 30 min at  
16 65° C with 2xSSC containing, 1% SDS followed by two additional  
17 washes for 30 min at 65° C with 0.1xSSC containing 1% SDS.

18       The filters were autoradiographed and duplicates were  
19 superimposed on each other to isolate PMA/chx-induced signals.  
20 Differential signals were plaque-purified by repeating the  
21 screening process. Insert cDNA was prepared and used for either  
22 Northern blot analysis or subcloning into plasmid vectors.

23       Of the twelve positive signals obtained from >10<sup>5</sup> pfu of  
24 the library three were found to be consistently positive. Two  
25 of the clones had inserts had sequences identical to the  
26 sequence of DNA that encodes human collagenase Type 1. The  
27 third clone, herein called edg-1 (1-Kb) contained a poly A  
28 tract at 3' end, a unique nucleotide sequence and hybridized  
29 to a 3.0 Kb PMA inducible mRNA species.

30       This 1 kb insert was used to rescreen two additional cDNA  
31 libraries-lambda gt10 and cDM8. The largest clone was 2.8  
32 kb. Further investigation and analysis was conducted using  
33 this clone, which is expressed at high levels (0.05%) in the  
34 HUVEC.

### EXAMPLE 2

The kinetics of edg RNA induction by PMA was studied by Northern blot analysis of HUVEC that were exposed to PMA for 0.5, 1, 2, and 4 hours (Figure 1 (A)).

In order to determine the characteristics of the rapid edg-1 induction, Northern blot analysis was performed with HUVEC that had been treated for 4 hours with PMA and chx, alone or in combination (Figure 2). As can be seen in Figure 2, the 3.0 KB mRNA edg transcript was induced independently by PMA and chx, but was superinduced in the presence of both.

### EXAMPLE 3

Chx was shown to exert the superinduction effect by stabilizing the *edg-1* transcript (Figure 3). HUVEC were stimulated for 4 hour with PMA and subsequently incubated with actinomycin D, in inhibitor of transcription both in the presence and absence of chx. As shown in Figure 3 steady-state levels of the *edg-1* mRNA declined to undetectable levels two hours after the addition of actinomycin D; whereas, chx prevented this decline.

### EXAMPLE 4

In order to ascertain at what level PMA induces *edg-1* mRNA, *edg 1* induction in the presence of actinomycin D was investigated. As shown in Figure 2, actinomycin D repressed the inductive effect of PMA, which suggests that PMA induces the transcription of the *edg-1* gene.

-18-

EXAMPLE 5Nuclear Run-On Transcription.

Nuclei ( $10^7$ ) were prepared from quiescent HUVEC untreated or treated with 20 ng/ml PMA for 2 hr. *In vitro* labeled, run-off transcripts were prepared by incubating the nuclei with 250  $\mu$ Ci of [ $\alpha$ - $^{32}$ P]-UTP (.6000 CI/mmol, Amersham), 10mM ATP, CTP, GTP, in the reaction buffer containing 20mM Tris-HCl, pH 7.9, 140mM KCl, 10mM MgCl<sub>2</sub> and 1mM dithiothreitol as described (Nevins, J., (1987) Meth. Enzymol. 152, 234-240).

The labeled RNA was purified (Winkles, J., supra.) and hybridized to nylon filters containing either 10  $\mu$ g of denatured plasmid edg-1 cDNA, 2  $\mu$ g of human fibronectin or 10  $\mu$ g of pBluescript (Stratagene). The hybridization and washing conditions were identical to those described for the differential hybridization.

Nuclei were prepared from untreated HUVEC or from HUVEC treated with PMA for 2 hours. Labeled run-on transcripts were obtained and hybridized to immobilized plasmid DNA containing the edg-1 insert and to a control plasmid containing fibronectin-encoding DNA or to a Bluescript plasmid (Figure 4). Edg-1 transcription was significantly induced in nuclei from the PMA treated HUVEC.

EXAMPLE 6DNA Sequence Analysis.

The structure of the edg-1 gene and gene product was elucidated by DNA sequencing of the 2.8 Kb cDNA clone.

Plasmid DNA for edg-1 (2.8Kb) was obtained by screening a cDNA library from HUVEC constructed in the vector, CDM8, which was a gift of Brian See, Harvard Medical School) with the (1.6Kb) insert obtained from the cDNA library in lambda

gt10, discussed in Example 1. Double-stranded sequence analysis was performed using the sequenase-2 enzyme (USBC), following the manufacturer's instructions. Successive primers were synthesized and used to sequence both strands of the cDNA clone. The DNA sequence was analyzed by the Intelligenetics Sequence Analysis program.

As shown in Figure 5, the complete nucleotide sequence of the edg-1 cDNA clone is 2774 bp long and, at nucleotide 251 from the 5' end, contains a consensus translation initiation sequence, which is followed by an open-reading frame (ORF) that encodes 380 amino acids. The ORF is followed by a 3', A/T-rich, 1.3 Kb untranslated region followed by a poly A tail. A/T rich sequence motifs in 3' untranslated regions have been implicated in conferring rapid RNA degradation of intermediate-early mRNAs. There are two consensus polyadenylation sites (AATAAA) at nucleotides 2590 and 2737, respectively. The edg-1 clone also contains about 250 bp of 5' untranslated region.

The deduced amino acid sequence contains a non-hydrophobic amino-terminal stretch of 46 amino acids, which contain two potential N-linked glycosylation sites at residues 29 and 35. This stretch is followed by seven alternating stretches of hydrophobic regions, each about 20 amino acid residues long. There are 8 hydrophilic regions. Each of the hydrophobic regions is flanked by hydrophilic regions of 7 to 19 amino acids, except for the region between the fifth and sixth transmembrane domain, which is 35 residues long and is rich in basic and dibasic residues. The last transmembrane domain is followed by a long, 66 amino acid, stretch of hydrophilic residues that include an abundance of serine and threonine residues.

EXAMPLE 7Reverse Transcriptase-Polymerase Chain Reaction Analysis

RNA from HUVEC was purified as described in Example 1. RNA from human saphenous vein smooth muscle cells, human foreskin fibroblasts, human epidermoid carcinoma cells (A431), human cervical carcinoma cells (HeLa), human melanocytes and total brain were the generous gift of Dr. Jeffrey Winkles of the American National Red Cross.

Total RNA (5 µg) from all the cultured cells and poly A<sup>+</sup>RNA (1 µg) from human brain (Clontech) was converted to cDNA by treatment with 200 units of MMLV reverse transcriptase (Bethesda Research Labs, MD) in 50 mM Tris-HCl, pH, 8.0, 1 mM dithiothreitol, 15 mM NaCl, 3 mM MgCl<sub>2</sub>, 1 unit RNasin (Promega), 0.2 µg of random hexamer primers, 0.8 mM dNTPs and incubated for 1 hour at 37° C. The reaction was terminated by heating at 95° C for 10 minutes and diluted to 1 ml with distilled water.

Enzymatic amplification was done on a 10 µl aliquot of the cDNA mix. PCR was performed in 50 mM Tris-HCl, pH, 8.0, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.2 mM dNTPs, 0.5 µg each of primers for edg-1 and 2.5 units of Taq DNA polymerase (Cetus, CA) (see, Saiki et al. (1988) Science 239, 487-491). The reaction mixture was heated at 94° C for 1 minute, annealed at 55° C for 2 minutes, and extended at 72° C for 3 minutes for 30 repetitive cycles. The primers used were as follows:

(1) 5'-TG TAC TGC AGA ATC TAC T-3' (sense) and 5'-T GCA GCC CAC ATC CAG CAG CA-3' (antisense) to amplify from nucleotide no. 909 to 1094, which spans the third cytosolic domain; and

(2) 5' AAG ACC TGT CAC ATC CTC TTC-3' (sense) and 5' ATG AAC CCT TTA GGA GCT TGA CAA-3' (antisense) to amplify from nucleotide no. 1100 to 1702, which spans the seventh transmembrane domain, the cytosolic tail and part of the 3'untranslated region.

-21-

1           When RNA from the various cultured human cell lines and  
2           from human brain was reverse transcribed and the cDNAs  
3           amplified using the oligonucleotides that are specific for the  
4           C-terminal domain (amino acids 266 to the termination codon  
5           and 309 bp of the 3' untranslated region, nucleotides 1100 to  
6           1702, see, e.g., Figures 5-7 and 9) an amplified product is  
7           the expected size, 600 bp., is observed (see Fig. 9 (A)) in  
8           RNA from all cell types and human brain. The intensity of the  
9           signal was most prominent in endothelial cells, but was  
10          present to a lesser extent in smooth muscle cells,  
11          fibroblasts, epidermoid cells, melanocytes, and brain tissue.

12          When the cDNAs were amplified with a pair of  
13          oligonucleotides that span the third intracellular loop (amino  
14          acids 220-282, nucleotides 909-1094), cell-specific bands were  
15          amplified (Figure 9 (B)). In smooth muscle cells, a major  
16          band at 0.7 Kb and minor bands at 0.9, 0.3, and 0.19 Kb were  
17          observed. In HeLa cells a very prominent band was observed  
18          at 0.3 Kb. The expected 0.19Kb amplification product was  
19          observed only in endothelial cells.

20          This result indicates that cDNAs derived from mRNAs that  
21          are related to, but not identical with, the *edg-1* transcript  
22          are present in different cell types and tissues. Because the  
23          third cytosolic loop has been identified in other G-protein-  
24          coupled receptors as the region that binds to the G-protein,  
25          the tissue specific transcripts differ in the region that  
26          encodes the portion of the receptor that couples with the G-  
27          protein and thereby modulates the cellular response of the  
28          particular cell type to the specific signal.

29          Since modifications will be apparent to those of skill  
30          in the art, it is intended that this invention be limited only  
31          by the scope of the appended claims.

1           We claim:

2           1. A purified DNA molecule that encodes a protein having  
3           the sequence of amino acids set forth in Figure 5.

4           2. The purified DNA molecule having the sequence of  
5           nucleotide bases set forth in Figure 5.

6           3. A purified protein that has substantially the same  
7           amino acid sequence as the sequence of amino acids set forth  
8           in Figure 5.

9           4. A purified DNA molecule that encodes the protein of  
10          claim 3.

11          5. A protein that includes regions that are  
12          substantially homologous with all or a portion of the protein  
13          of Figure 5, wherein said portion consists of the amino acids  
14          that comprise the transmembrane domains of the protein of  
15          Figure 5.

16          6. A protein selected from the group consisting of the  
17          edg-1-G-coupled-protein receptor family of proteins.

18          7. The protein of claim 6, that is expressed in a cell  
19          or tissue selected from the group consisting of smooth muscle  
20          cells, fibroblasts, cultured immortal human cell lines,  
21          epidermoid carcinoma cells, melanocytes, brain tissue and  
22          differentiating endothelial cells.

23          8. An isolated DNA molecule that encodes the protein of  
24          claim 7.



1/13

# FIG. 1

(A) 0 .5 1 2 4 hrs.



(B)



2/13

**FIG. 2**

PMA	-	-	+	+	+
CHX	-	+	-	+	-
Act D	-	-	-	-	+

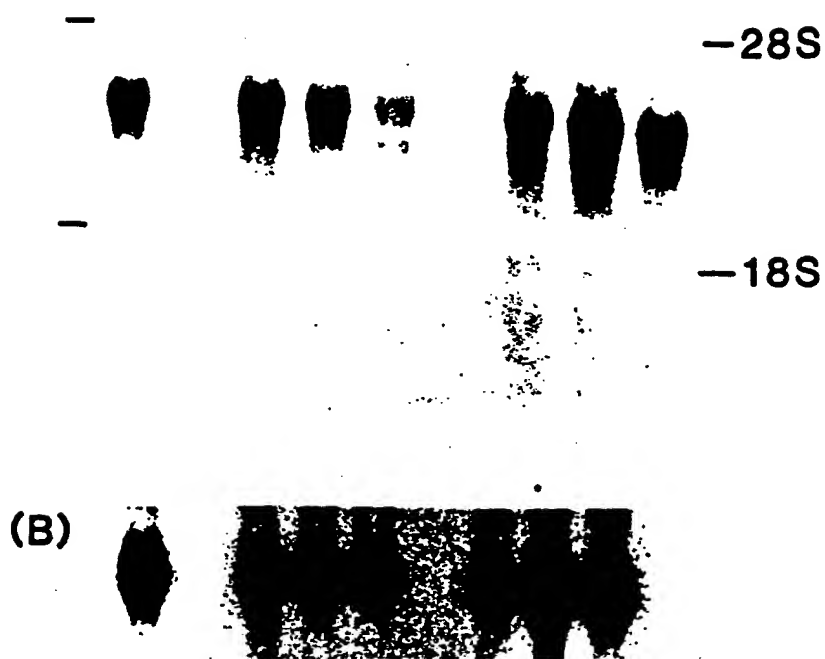
**(A)****(B)****SUBSTITUTE SHEET**

3/13

**FIG. 3**

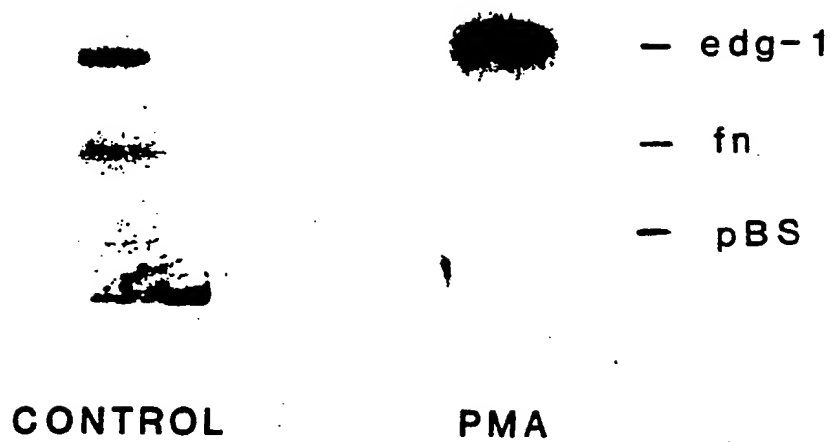
Act D	—	+	+	+	+	+	+
CHX	—	—	—	—	+	+	+
	0'	15'	30'	120'	15'	30'	120'

(A)

**SUBSTITUTE SHEET**

4/13

**FIG. 4**



10	20	30	40	50	60	70
TCTAAAGGTC	GGGGGCAGCA	GCAAGATCGG	AAGCGAGCCG	TACAGATCCC	GGGCTCTCCG	AACGCAACTT
80	90	100	110	120	130	140
CGCCCTGCTT	GAGCGAGGCT	GCGGTTTCCG	AGGCCCTCTC	CAGCCAAGGA	AAAGCTACAC	AAAAAGCCTG
150	160	170	180	190	200	210
GATCACTCAT	CGAACCACCC	CTGAAGCCAG	TGAAGGCTCT	CTCGCCTCGC	CCTCTAGCGT	TCGTCTGGAG
220	230	240	250	259	268	
TAGCGCCACC	CCGGCTTCCT	GGGGACACAG	GGTTGGCACC	ATG GGG CCC	ACC AGC GTC	CCG
277	286	295	304	313	322	
CTG GTC AAG	GCC CAC CGC	AGC TCG GTC	TCT TCT GAC	TAC GTC AAC	TAT GAT ATC	ATC
Leu Val Lys	Ala His Arg	Ser Ser Val Ser	Asp Asp Tyr Val	Asn Tyr Asp	Ile Ile Ile	Ile
331	340	349	358	367	376	5/13
GTC CGG CAT	TAC AAC TAC	ACG GGA AAG	CTG AAT ATC	AGC GCG GAC	AAG GAG AAC	
Val Arg His	Tyr Asn Tyr	Thr Gly Lys	Leu Asn Ile	Ser Ala Asp	Lys Glu Asn	
385	394	403	412	421	430	
AGC ATT AAA	CTG ACC TCG	GTG GTG TTC	ATT CTC ATC	TGC TGC TTT	ATC ATC CTG	
Ser Ile Lys	Leu Thr Ser	Val Val Phe	Ile Leu Ile	Cys Cys Phe	Ile Ile Leu	
439	448	457	466	475	484	
GAG AAC ATC	TTT GTC TTG	CTG ACC ATT	TGG AAA ACC	AAG AAA TTC	CAC CGA CCC	
Glu Asn Ile	Phe Val Leu	Leu Thr Ile Trp	Lys Thr Lys	Lys Phe His	Arg Pro	
493	502	511	520	529	538	
ATG TAC TAT	TTT ATT GGC	AAT CTG	GCC CTC TCA	GAC CTG	TTG GCA	GTA GCC
MET Tyr Tyr	Phe Ile Gly	Asn Leu Ala	Leu Ser Asp	Leu Leu Ala	Gly Val Ala	

FIG. 5A

547	ACA	GCT	AAC	CTG	CTC	TTG	TCT	565	GGG	GCC	ACC	ACC	TAC	AAG	583	CTC	ACT	CCC	592	GCC
	Tyr	Thr	Ala	Asn	Leu	Leu	Leu	Ser	Gly	Ala	Thr	Thr	Tyr	Lys		Leu	Thr	Pro		Ala
601	TGG	TTT	CTG	CGG	GAA	GGG	AGT	619	ATG	TTT	GTG	GCC	CTG	TCA	637	GCC	TCC	GTG	646	TTC
	Gln	Trp	Phe	Leu	Arg	Glu	Gly	Ser	MET	Phe	Val	Ala	Leu	Ser		Ala	Ser	Val		Phe
655	CTC	CTC	GCC	ATC	GCC	ATT	GAG	673	CGC	TAT	ATC	ACA	ATG	CTG	691	AAA	ATG	AAA	700	CTC
	Ser	Leu	Leu	Ala	Ile	Ala	Ile	Glu	Arg	Tyr	Ile	Thr	MET	Leu		Lys	MET	Lys		Leu
709	AAC	GGG	AGC	AAT	AAC	TTC	CGC	727	CTC	TTC	CTG	CTA	ATC	AGC	745	GCC	TGC	TGG	754	GTC
	His	Asn	Gly	Ser	Asn	Asn	Phe	Arg	Leu	Phe	Leu	Leu	Ile	Ser		Ala	Cys	Trp		Val
763	TCC	CTC	ATC	CTG	GGT	GGC	CTG	781	CCT	ATC	ATG	GGC	TGG	AAC	799	TGC	ATC	AGT	808	GCG
	Ile	Ser	Leu	Ile	Leu	Gly	Gly	Leu	Pro	Ile	MET	Gly	Trp	Asn		Cys	Ile	Ser		Ala
817	TCC	AGC	TGC	TCC	ACC	GTG	CTG	835	CCG	CTC	TAC	CAC	AAG	CAC	853	TAT	ATC	CTC	862	TTC
	Leu	Ser	Ser	Cys	Ser	Thr	Val	Leu	Pro	Leu	Tyr	His	Lys	His		Tyr	Ile	Leu		Phe
871	ACC	ACG	GTC	TTC	ACT	CTG	CTT	889	CTG	CTC	TCC	ATC	GTC	ATT	907	CTG	TAC	TGC	916	AGA
	Cys	Thr	Thr	Val	Phe	Thr	Leu	Leu	Leu	Leu	Ser	Ile	Val	Ile		Leu	Tyr	Cys		Arg

FIG. 5B

0

925	TAC	TCC	TTG	GTC	AGG	ACT	CGG	AGC	CGC	CGC	952	961	970
	Ile	Tyr	Ser	Leu	Val	Arg	Thr	Arg	Ser	Arg	Leu	Thr	Phe
													Arg
													Lys
													Asn
													Ile
979	AAG	GCC	AGC	AGC	AGC	TCT	GAG	AAT	GTG	GCG	1006	1015	1024
	TCC	AAG	GCC	AGC	AGC	TCT	GAG	AAT	GTG	GCG	CTG	CTC	AAG
	Ser	Lys	Ala	Ser	Arg	Ser	Glu	Asn	Val	Ala	Leu	Lys	ACC
													GTA
													ATT
													ATC
1033	CTG	AGC	GTC	TTC	ATC	GCC	TGC	TGG	GCA	CCG	1060	1069	1078
	Val	Leu	Ser	Val	Phe	Ile	Ala	Cys	Trp	Ala	Pro	Leu	CTG
													CTC
													CTG
													Leu
1087	GTG	GGC	TGC	AAG	GTG	AAG	ACC	TGT	GAC	ATC	1114	1123	1132
	GAT	GTG	GGC	TGC	AAG	ACC	TGT	GAC	ATC	CTC	CTC	TTC	AGA
	Asp	Val	Gly	Cys	Lys	Val	Lys	Thr	Cys	Asp	Ile	Leu	Phe
													Arg
													Ala
													Glu
													Tyr
													Phe
1141	GTG	TTA	GCT	GTG	CTC	AAC	TCC	GGC	ACC	AAC	1168	1177	1186
	CTG	GTG	TTA	GCT	GTG	CTC	AAC	TCC	GGC	ACC	CCC	ATC	ATT
	Leu	Val	Leu	Ala	Val	Leu	Asn	Ser	Gly	Thr	Asn	Pro	Ile
													Ile
													Tyr
													Thr
1195	AAG	GAG	ATG	CGT	CGG	GCC	TTC	ATC	CGG	ATC	1222	1231	1240
	AAC	AAG	GAG	ATG	CGT	CGG	GCC	TTC	ATC	CGG	ATC	TCC	TGC
	Asn	Lys	Glu	MET	Arg	Arg	Ala	Phe	Ile	Arg	Ile	MET	Ser
													Cys
													Lys
													Cys
													Pro
1249	GGA	GAC	TCT	GCT	GGC	AAA	TTC	AAG	CGA	CCC	1276	1285	1294
	AGC	GGA	GAC	TCT	GCT	GGC	AAA	TTC	AAG	CGA	CCC	ATC	ATC
	Ser	Gly	Asp	Ser	Ala	Gly	Lys	Phe	Lys	Arg	Pro	Ile	Ile
													Ala
													Gly
													MET
													Glu
													Phe

FIG. 5C

8/13

1303	1312	1321	1330	1339	1348
AGC CGC AGC AAA TCG GAC AAT TCC TCC CAC CCC CAG AAA GAC GAA GGG GAC AAC					
Ser Arg Ser Lys Ser Asp Asn Ser Ser His Pro Gln Lys Asp Glu Gly Asp Asn					
1357	1366	1375	1384	1393	1406
CCA GAG ACC ATT ATG TCT TCT GGA AAC GTC AAC TCT TCT TCC TAG AACTGGAAGC					
Pro Glu Thr Ile MET Ser Ser Gly Asn Val Asn Ser Ser Ser					
1416	1426	1436	1446	1456	1466
TGTCCACCCA CCGGAAGCGC TCTTTACTTG GTCGCTGGCC ACCCCAGTGT TTGGAAAAA ATCTCTGGGC					
1486	1496	1506	1516	1526	1536
TTCGACTGCT GCCAGGGAGG AGCTGCTGCA AGCCAGAGGG AGGAAGGGG AGAATACGAA CAGCCTGGTG					
1556	1566	1576	1586	1596	1606
GTGTCGGGTG TTGGTGGTA GAGTTAGTTC CTGTGAACAA TGCACCTGGA AGGGTGAGA TCAGGTCCCG					
1626	1636	1646	1656	1666	1676
GCCTGGAATA TATATTCTAC CCCCCTGGAG CTTTGATTTT GCACTGAGCC AAAGGTCTAG CATTGTCAAG					
1696	1706	1716	1726	1736	1746
CTCCTAAAGG GTTCATTGG CCCCCTCCTCA AAGACTAATG TCCCCATGTG AAAGCGTCTC TTTGTCTGGA					
1766	1776	1786	1796	1806	1816
GCTTTGAGGA GATGTTTCC TTCACCTTAG TTCAAAACCC AAGTGAGTGT GTGCACTTCT GCTTCTTAG					
1836	1846	1856	1866	1876	1886
GGATGCCCTG TACATCCCAC ACCCCACCCT CCCTTCCCTT CATACCCCTC CTCAACGTTT TTTTACTTTA					
1906	1916	1926	1936	1946	1956
TACTTTAACT ACCTGAGAGT TATCAGAGCT GGGGTTGTGG AATGATCCGAT CATCTATAGC AAATAGGCTA					
1976	1986	1996	2006	2016	2026
TGTTGAGTAC GTAGGCTGTG GGAAGATGAA GATGGTTTGG AGGTGTAAAA CAAIGTCCTT CGCTGAGGCC					

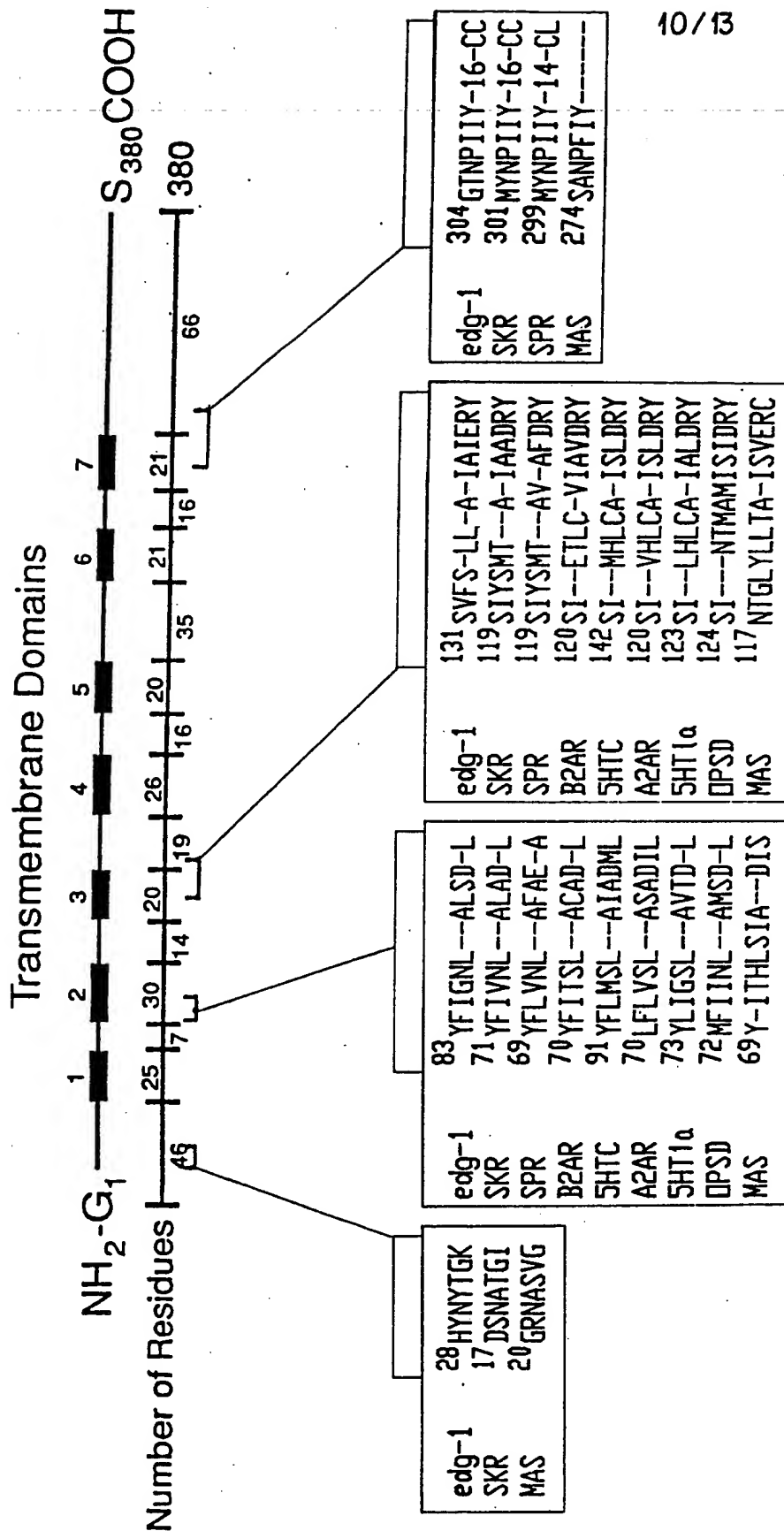
FIG. 5D



2046	2056	2066	2076	2086	2096	2106
AAAGTTTCCA	TGTAAGCGGG	ATCCGTTTTT	TGGAATTGG	TTGAAGTCAC	TTTGATTCT	TTAAAAACA
2116	2126	2136	2146	2156	2166	2176
TCTTTTCAAT	GAAATGTGTT	ACCATTTTCA	ATCCATTGAA	GCCGAAATCT	GCATAAGGAA	GCCCACTTTA
2186	2196	2206	2216	2226	2236	2246
TCTAAATGAT	ATTAGCCAGG	ATCCTTGGTG	TCCTAGGAGA	AACAGACAAG	CAAAACAAAG	TGAAAAACCGA
2256	2266	2276	2286	2296	2306	2316
ATGGATTAA	TTTTGCAAAC	CAAGGGAGAT	TTCTTAGCAA	ATGAGTCTAA	CAAATATGAC	ATCCGTCTTT
2326	2336	2346	2356	2366	2376	2386
CCCACTTTTG	TTGATGTTTA	TTTCAGAAATC	TTGTGTGATT	CATTTCAGC	AACAACATGT	TGTATTTTGT
2396	2406	2416	2426	2436	2446	2456
TGTGTTAAAA	GTAATTTTCT	TGATTTTGA	ATGTATTTGT	TTCAGGAAGA	AGTCATTTTA	TGGATTTTTC
2466	2476	2486	2496	2506	2516	2526
TAACCCCGTG	TAACTTTTCT	AGAATCCACC	CTCTTGTC	CTTAAGCATT	ACTTTAACTG	GTAGGGAACG
2536	2546	2556	2566	2576	2586	2596
CCAGAACTTT	TAAGTCCAGC	TATTCATTAG	ATAGTAATTG	AAGATATGTA	TAAATATTAC	AAAGAAATAA
2606	2616	2626	2636	2646	2656	2666
AATATATTAC	TGTCTCTTTA	GTATGGTTTT	CAGTGCAATT	AAACCGAGAG	ATGTCCTGTT	TTTTTAAAAA
2676	2686	2696	2706	2716	2726	2736
GAATAGTATT	TAAATAGGTT	CTGACTTTTG	TGGATCATTT	TGCACATAGC	TTTATCAACT	TTTAAACATT
2746	2756	2766				
AAATAAACTGA	TTTTTTTAA	GAAAAAAA	AAAAAAAG			
=====						

9 / 13

FIG. 5E



10/13

FIG. 6

11/13

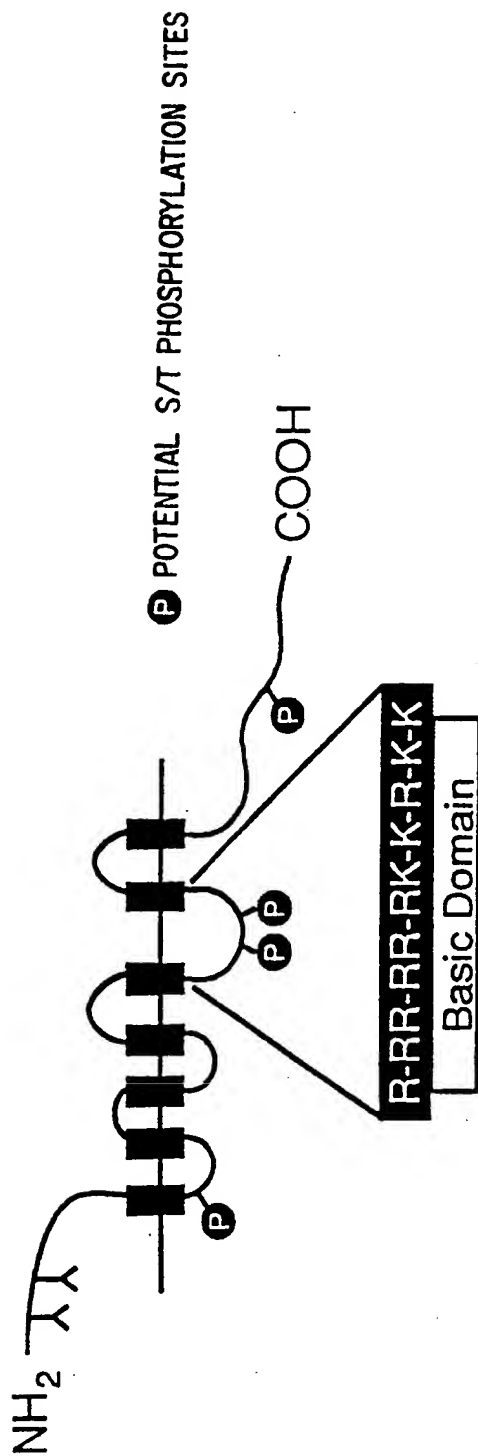


FIG. 7

12 / 13

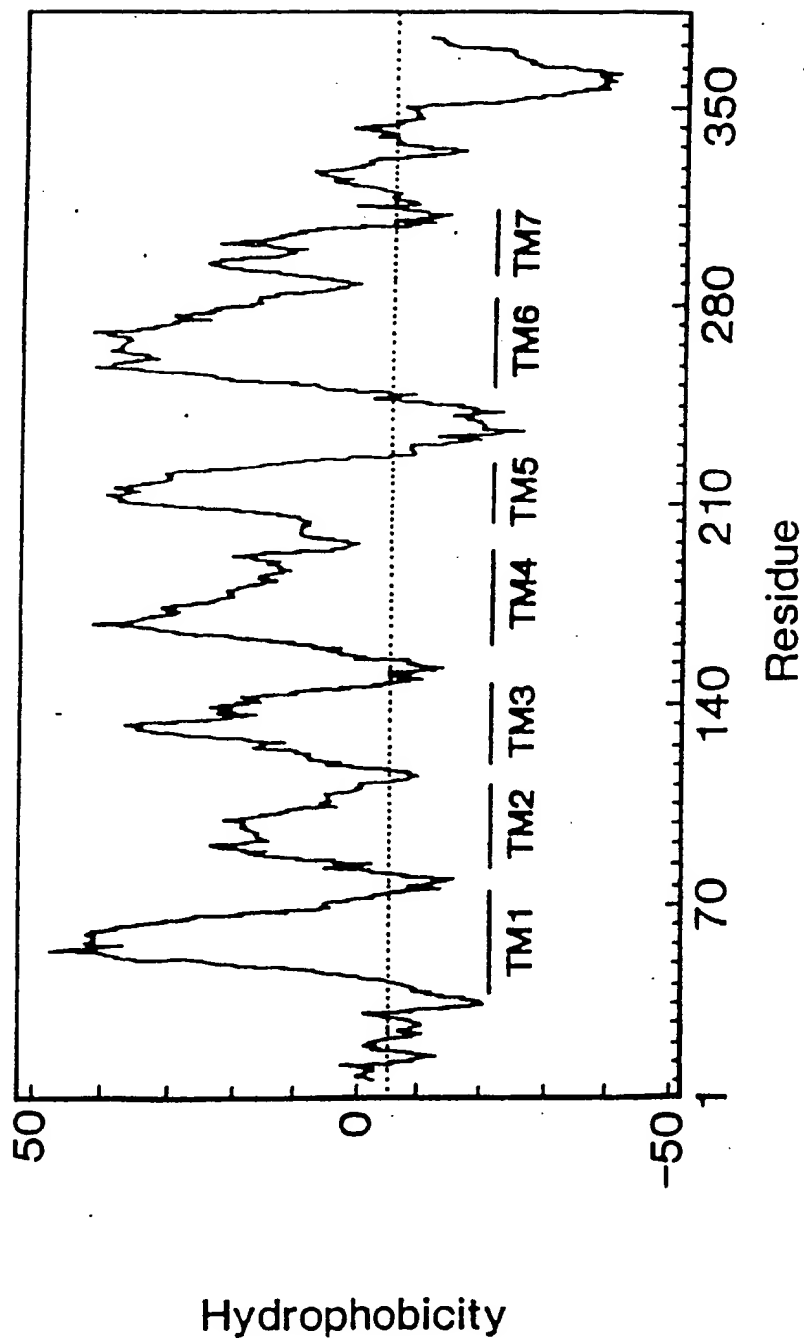


FIG. 8

SUBSTITUTE SHEET

13 / 13

FIG. 9B

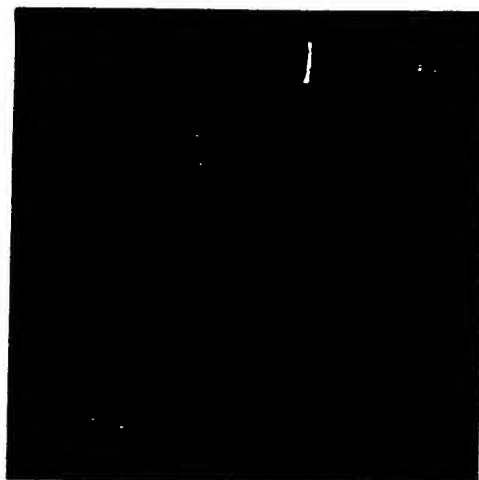
S F A H M B E



^ ^ ^ ^ ^ ^

FIG. 9A

S F A H M B E



^ ^ ^ ^ ^ ^

## INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/02344

I. CLASSIFICATION OF SUBJECT MATTER (If several classification symbols apply, indicate all *)		
According to International Patent Classification (IPC) or to both National Classification and IPC		
IPC(5): C12N 15/12; C07K 15/06, 15/14		
U.S.CL.: 536/27, 530/350,395		
II. FIELDS SEARCHED		
Minimum Documentation Searched *		
Classification System	Classification Symbols	
U.S.C1.	536/27; 530/350,395.	
Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched *		
APS AND DIALOG Files 357,155,WPI,72,35,5 and 399 searched for edg type receptor proteins and sequences.		
III. DOCUMENTS CONSIDERED TO BE RELEVANT *		
Category *	Citation of Document, ** with indication, where appropriate, of the relevant passages **	Relevant to Claim No. **
X,P	Journal of Biological Chemistry, vol. 265, No. 16, issued 05 June 1990. Hla et al.. "An abundant transcript induced in differentiating human endothelial cells encodes a polypeptide with structural similarities to G-protein- coupled receptors", pages 9308-9313. See whole publication, especially the abstract, p. 9309 and 9311.	1-8
X	Science, vol. 241, issued 16 September 1988. Klein et al., "A chemoattractant receptor controls development in <u>Dictyostelium</u> <u>discoideum</u> ", pages 1467-1472. See whole publications, especially Figure 8 on p. 1472.	3-8
A	Science, vol. 245, issued 08 September 1989; Devreotes. " <u>Dictyostelium discoideum</u> : a model system for cell-cell interactions in development". pages 1054-1058. See whole publication.	1-8
<p>* Special categories of cited documents: **</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claims or which is cited to establish the publication date of another citation or other special reason (is specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such docu- ments, such combination being obvious to a person skilled in the art</p> <p>"Z" document member of the same patent family</p>		
IV. CERTIFICATION		
Date of the Actual Completion of the International Search		Date of Mailing of the International Search Report
02 July 1991		26 JUL 1991
International Searching Authority		Signature of the International Searching Authority
ISA/US		Keith C. Furman KEITH C. FURMAN